Using prospect theory to investigate the low value of travel time for small time changes

Hjorth, Katrine; Fosgerau, Mogens

Publication date:
2011

Document Version
Peer reviewed version

Using prospect theory to investigate the low value of travel time for small time changes *

Katrine Hjorth
Department of Transport, Technical University of Denmark.
Department of Economics, University of Copenhagen.

Mogens Fosgerau
Department of Transport, Technical University of Denmark.
Centre for Transport Studies, Royal Institute of Technology.

Paper submitted to the International Choice Modelling Conference 2011

Abstract

A common finding in stated preference studies that measure the value of travel time (VTT), is that the measured per-minute VTT increases with the size of the time change considered, in conflict with standard neoclassical theory. The current paper tests prospect theory as a possible explanation: More specifically, whether the phenomenon is generated by preferences being reference-dependent and exhibiting diminishing sensitivity for gains and losses, with a stronger degree of diminishing sensitivity for money than for travel time.

We use stated preference data with trade-offs between travel time and money that provide identification of the degrees of diminishing sensitivity for time and money gains and losses, thus enabling us to test and potentially falsify the prospect theory explanation. We apply a discrete choice model, in which choice depends on a reference-free value of travel time and reference-dependent value functions for time and money, allowing for loss aversion and different degrees of diminishing sensitivity for gains and losses. We use semi-parametric local logit estimates of the equi-probability curves in the data to test the model's appropriateness, and estimate its parameters using a mixed logit approach. Our results support the prospect theory explanation.

*The authors would like to thank The Institute of Transport Economics, Oslo, for use of their data, and The Danish Social Science Research Council for financial support.

1 Introduction

An often encountered phenomenon in stated preference (SP) studies that measure the value of travel time (VTT), is that the measured per-minute VTT increases with the size of the time change considered, in conflict with standard neoclassical theory (Gunn, 2001; Hultkrantz and Mortazavi, 2001; Mackie et al., 2001, 2003; Fosgerau et al., 2007). The effect is large enough to be of considerable economic significance (Mackie et al., 2003; Fosgerau et al., 2007), and problematic because it is inappropriate for evaluations of transport projects to apply a lower unit VTT for small time changes: This would cause evaluations to depend in an illogical way on whether the project was evaluated as a whole or as a series of smaller projects each resulting in smaller time savings (Fosgerau et al., 2007).

Several explanations to the phenomenon have been proposed (Mackie et al., 2003; Cantillo et al., 2006), but so far it remains a puzzle. Recently, de Borger and Fosgerau (2008) suggested prospect theory as a possible explanation, arguing that the phenomenon could be generated by preferences being reference-dependent and exhibiting diminishing sensitivity for gains and losses, with a stronger degree of diminishing sensitivity for money than for travel time. This explanation is supported by the fact that stated preference studies measuring the VTT until quite recently did not take reference-dependence into account.[1]

For the explanation to be valid, two conditions must hold: First, the reference-dependent model underlying the analysis in de Borger and Fosgerau (2008) must be an adequate description of the behaviour observed in the SP surveys. Second, the observed preferences should exhibit stronger diminishing sensitivity for money than for travel time. De Borger and Fosgerau (2008) provide empirical support for the latter condition, but only partly for the former, because they lack the data to separately identify the degrees of diminishing sensitivity for travel time and cost. The current paper extends their analysis, using data that provide better identification, and thus presents an empirical test with potential to falsify the prospect theory explanation.

Usually, the VTT is measured from SP data where respondents make choices between travel alternatives that differ with respect to travel time and cost. A common experimental setup is to use binary choices between a fast and expensive travel alternative and a slower and cheaper one. In recent studies, using electronic questionnaires, the time and cost attributes of the alternatives are varied around individual-specific reference values, corresponding to the normal or most recently experienced travel time and cost of the journey of interest (Burge et al., 2004; Fosgerau et al., 2007; de Jong et al., 2007; Ramjerdi et al., 2010). Table 1 presents four types of choices often applied in such VTT studies, using the following notation: Let t_1, t_2, c_1, c_2 be the travel time and cost attributes of the two alternatives, respectively, normalised by subtracting the reference values, such that negative values correspond to gains (faster or cheaper than reference) and positive values to losses (slower or more expensive than reference). Assume alternatives are sorted such that $t_1 < t_2$ and $c_1 > c_2$, and define $\Delta t := t_2 - t_1$ and $\Delta c := c_1 - c_2$. We use the nota-

[1] Descriptive behavioural theories as prospect theory and rank-dependent utility theory have only recently been applied in travel behaviour research (see, e.g. Van de Kaa, 2008; Avineri and Bovy, 2008). To our knowledge, Van de Kaa (2005) was one of the first to argue that VTT studies should control for reference-dependence, preceded by a discussion of the gap between willingness-to-pay and willingness-to-accept in such studies. Recent VTT studies have allowed for reference-dependence in the form of loss aversion, whereas diminishing sensitivity for gains and losses is generally not accommodated.

tion from de Borger and Fosgerau (2008) and label the choice types WTP (willingness-to-pay), WTA (willingness-to-accept), EG (equivalent gain), and EL (equivalent loss). The choices are reference-based in the sense that they always have one time attribute equal to the reference time (i.e. $t_1 = 0$ or $t_2 = 0$) and one cost attribute equal to the reference cost (i.e. $c_1 = 0$ or $c_2 = 0$). [2]

In such a setting, if the reference values represent the respondent's perception of the normal travel time and cost, prospect theory suggests that the indirectly observed preferences may be reference-dependent (Kahneman and Tversky, 1979; Tversky and Kahneman, 1991). In prospect theory, preferences are defined in terms of value functions, which have three general characteristics: *Reference-dependence:* the carriers of value are gains and losses relative to a reference point; *Loss aversion:* losses are valued more heavily than gains; *Diminishing sensitivity:* the marginal value of both gains and losses decreases with their size.

De Borger and Fosgerau (2008) analyse data of the type presented in Table 1, using a choice model with reference-dependent preferences for travel time and money, based on prospect theory. They use a flexible functional form for the value functions for time and money, which permits the characteristics of prospect theory, but is more general. However, the authors are unable to identify value function curvature empirically (they can only identify the ratio of time and money curvature parameters) because their data only contain reference-based choice situations.

This paper extends their analysis by also using two types of non-reference-based choices, shown in Table 2. Here, both time attributes are different from the reference time. Using the modelling framework from de Borger and Fosgerau (2008), we formulate a discrete choice model, in which choice depends on the reference-free value of travel time and the value functions for time and cost. We test this parametric model by comparing its predicted equi-probability curves to those of the data, estimated using a semi-parametric local logit estimator (Fan et al., 1995; Fosgerau, 2007). Based on this test, we conclude that our data do not reject the parametric model.

The value functions are estimated from our parametric model using mixed logit estimation, and the results are consistent with prospect theory. In general, the value functions exhibit loss aversion for both travel time and cost, the value function for cost exhibits diminishing sensitivity for both gains and losses, and the value function for time exhibits constant sensitivity for both gains and losses. We find that the value function for cost "bends" more than the value function for time, i.e. there is stronger diminishing sensitivity for money than for travel time. Our results thus support prospect theory as an explanation of the phenomenon that VTT increases with the size of the time change.

The paper is organised as follows. Section 2 presents the model, section 3 our data, section 4 our analysis, and section 5 concludes.

[2]These choice types are applied in the national British (1994-96), Dutch (1988, 1997-98, 2007-), Danish (2004-2007) and Norwegian (2009) VTT studies (Burge et al., 2004; Fosgerau et al., 2007; de Jong et al., 2007; Ramjerdi et al., 2010). In addition, the Dutch and Norwegian studies included choices that were not reference-based. The national Swedish (1994) VTT study used a variation of the WTA and WTP choices (Burge et al., 2004).

Table 1: Reference-based choice types

	Fast alternative		Slow alternative	
Choice type	t_1	c_1	t_2	c_2
WTP	$-\Delta t$	Δc	0	0
WTA	0	0	Δt	$-\Delta c$
EL	0	Δc	Δt	0
EG	$-\Delta t$	0	0	$-\Delta c$

Note: $\Delta t, \Delta c > 0$ denote the time and cost differences between alternatives.

Table 2: Non-reference-based (nrb) choice types

	Fast alternative		Slow alternative	
Choice type	t_1	c_1	t_2	c_2
EL-nrb	t'	Δc	$t' + \Delta t$	0
EG-nrb	$-t' - \Delta t$	0	$-t'$	$-\Delta c$

Note: $\Delta t, \Delta c > 0$ denote the time and cost differences between alternatives. $t' > 0$ denotes the shift off the reference.

2 Model

2.1 Parametric model

Our setting is similar to the one in de Borger and Fosgerau (2008): We consider binary choices between two travel alternatives that differ with respect to travel time and cost, such that one alternative is faster but more expensive than the other. Individuals have a reference travel time t_0 and a reference cost c_0, representing their normal state. As above, t_1, t_2, c_1, c_2 denote the travel time and cost attributes of the two alternatives, respectively, normalised by subtracting the reference values, and alternatives are sorted such that $t_1 < t_2$ and $c_1 > c_2$.

Assume we observe the six different types of choices given in Tables 1 and 2. We assume that individuals prefer the slow alternative (alternative 2) whenever [3]

$$wv_t(t_1) + v_c(c_1) < wv_t(t_2) + v_c(c_2), \tag{1}$$

where w is a reference-free value of travel time (the absolute value of the reference-free marginal rate of substitution between travel time and money), which varies randomly in the population, and v_t, v_c are value functions for travel time and cost, that measure the values the individuals assign to the time and cost attributes.[4] As de Borger and Fosgerau (2008), we assume the value

[3]Unlike de Borger and Fosgerau (2008), we do not put w inside the value function for travel time.

[4]The term ”value function” stems from prospect theory (Kahneman and Tversky, 1979).

functions have the following form [5]

$$v_t(t) = -|t|^{1-\beta_t+\gamma_t S(t)} S(t) e^{\eta_t S(t)}, \tag{2}$$

$$v_c(c) = -|c|^{1-\beta_c+\gamma_c S(c)} S(c) e^{\eta_c S(c)}. \tag{3}$$

$S(\cdot)$ is the sign function, which takes the values 1, 0, and -1, when its argument is positive, zero, and negative, respectively. The parameters η, β, and γ determine the slope and curvature of the value functions. Equations (2) and (3) are flexible formulations, allowing for several possible shapes. We require the value functions to be decreasing, corresponding to $\beta - 1 < \gamma < 1 - \beta$. The value functions exhibit diminishing sensitivity to gains if $-\beta < \gamma$, and to losses if $\gamma < \beta$. If $\gamma = 0$, loss aversion is equivalent to $\eta > 0$. If $\gamma > 0$, the value function exhibits loss aversion for all values numerically larger than $\exp(-\eta/\gamma)$, while if $\gamma < 0$, we have loss aversion for all values numerically smaller than $\exp(-\eta/\gamma)$.

For the choice types in our data, it is always the case that

- either $c_1 = 0$ or $c_2 = 0$,

and

- either $t_1 = 0$ or $t_2 = 0$ or $S(t_1) = S(t_2)$.

Applying this with the value functions in equations (2) and (3), and taking logs, we see that eq. (1) is equivalent to

$$\begin{aligned} \log w < {} & \eta_c S(c_1+c_2) - \eta_t S(t_1+t_2) \\ & + \log\left[S(c_1+c_2)(|c_1|^{1-\beta_c+\gamma_c S(c_1+c_2)} - |c_2|^{1-\beta_c+\gamma_c S(c_1+c_2)})\right] \\ & - \log\left[S(t_1+t_2)(|t_2|^{1-\beta_t+\gamma_t S(t_1+t_2)} - |t_1|^{1-\beta_t+\gamma_t S(t_1+t_2)})\right]. \end{aligned} \tag{4}$$

Note that the terms in square brackets are always positive, so that the logarithms are well-defined. Let $y = 1_{\{\text{slow alt chosen}\}}$, i.e. y takes the value 1 when the slow alternative is chosen, and the value 0 otherwise. To take into account that individuals may make errors when comparing alternatives in the questionnaire, we do not assume that individuals choose the slow alternative whenever eq. (4) holds, but only that people do not deviate systematically from this rule. More specifically, we assume that

$$\begin{aligned} & y = 1 \\ & \Updownarrow \\ & \log w + \varepsilon < \eta_c S(c_1+c_2) - \eta_t S(t_1+t_2) \\ & \qquad + \log\left[S(c_1+c_2)(|c_1|^{1-\beta_c+\gamma_c S(c_1+c_2)} - |c_2|^{1-\beta_c+\gamma_c S(c_1+c_2)})\right] \\ & \qquad - \log\left[S(t_1+t_2)(|t_2|^{1-\beta_t+\gamma_t S(t_1+t_2)} - |t_1|^{1-\beta_t+\gamma_t S(t_1+t_2)})\right], \end{aligned} \tag{5}$$

[5] This is a two-part power function with separate slopes and exponents for gains and losses, as is often applied in studies based on prospect theory, though parameterized slightly differently. The power functional form has been criticized, because the measured degree of loss aversion depends on the scaling of the attributes (see e.g. Wakker, 2010); it has however, in the few comparisons available, been found to have empirical support in terms of better goodness-of-fit (Stott, 2006).

Table 3: Slopes and intercepts of equi-probability curves with prob. p in $(\log\Delta t, \log\Delta c)$-space.

Choice type	Slope	Intercept
WTP	$\frac{1-\beta_t-\gamma_t}{1-\beta_c+\gamma_c}$	$\frac{F^{-1}(p)-\eta_c-\eta_t}{1-\beta_c+\gamma_c}$
WTA	$\frac{1-\beta_t+\gamma_t}{1-\beta_c-\gamma_c}$	$\frac{F^{-1}(p)+\eta_c+\eta_t}{1-\beta_c-\gamma_c}$
EL	$\frac{1-\beta_t+\gamma_t}{1-\beta_c+\gamma_c}$	$\frac{F^{-1}(p)-\eta_c+\eta_t}{1-\beta_c+\gamma_c}$
EG	$\frac{1-\beta_t-\gamma_t}{1-\beta_c-\gamma_c}$	$\frac{F^{-1}(p)+\eta_c-\eta_t}{1-\beta_c-\gamma_c}$

where ε is a symmetric random error with mean zero, independently and identically distributed across individuals and choices.

2.2 Equi-probability curves for WTP, WTA, EG, and EL choices

For the choice types WTP, WTA, EG, and EL, we always have that $t_1 = 0$ or $t_2 = 0$, which implies that the probability of choosing the slow alternative can be written as a function of Δt, Δc, and F, the cumulative distribution function (CDF) of $\log w + \varepsilon$. Assume that $\log w + \varepsilon$ is a continuous random variable, such that F has an inverse. Then for WTP choices, where $t_2 = 0$ and $c_2 = 0$, we have that

$$\begin{aligned} p &= P(y = 1|\Delta t, \Delta c) \\ &= F\left(\eta_c + \eta_t + (1-\beta_c+\gamma_c)\log\Delta c - (1-\beta_t-\gamma_t)\log\Delta t\right) \end{aligned}$$

$$\Updownarrow$$

$$\log\Delta c = \frac{F^{-1}(p)-\eta_c-\eta_t}{1-\beta_c+\gamma_c} + \frac{1-\beta_t-\gamma_t}{1-\beta_c+\gamma_c}\log\Delta t \tag{6}$$

Hence the equi-probability curves in $(\log\Delta t, \log\Delta c)$-space, i.e. the sets $\{(\log\Delta t, \log\Delta c) \in \mathbb{R}^2 | P(y = 1|\Delta t, \Delta c) = p\}$ for different values of $p \in]0,1[$, are parallel straight lines. This is also the case for WTA, EG, and EL choices. Table 3 lists the slopes and intercepts for all four choice types.

Assume that the value functions are decreasing, i.e. that $\beta_t - 1 < \gamma_t < 1 - \beta_t$ and $\beta_c - 1 < \gamma_c < 1 - \beta_c$. This implies that the equi-probability curves have positive slopes, cf. Table 3. If $\gamma_t > 0$, the equi-probability curves will be steeper for EL than WTP choices, and steeper for WTA than EG choices. If $\gamma_c > 0$, the curves are steeper for EG than WTP choices, and steeper for WTA than EL choices. Moreover, loss aversion in the travel time dimension is equivalent to the equi-probability curve for EL being above that for WTP for a given value of p, and to the equi-probability curve for WTA being above that for EG.

2.3 Consequences of ignoring reference-dependence: A positive relation between the VTT and Δt

Suppose we could observe choices without any measurement error, and that everybody in the population had identical preferences and behaved according to equations (1), (2), and (3). What would happen if we tried to measure the VTT from standard data as the choice types in Table 1, but did not take reference-dependence into account? Let $\Delta t > 0$ denote a given time change, and consider the elicitation measure $WTP(\Delta t)$, defined as the cost change $\Delta c > 0$ that would make respondents indifferent between the two alternatives in a WTP choice. This measure is one possible estimate of the VTT. From equations (1), (2), and (3), it follows that (cf. the results in de Borger and Fosgerau, 2008):

$$WTP(\Delta t) = \left(we^{-\eta_t-\eta_c}\Delta t^{1-\beta_t-\gamma_t}\right)^{1/(1-\beta_c+\gamma_c)}.$$

Defining $WTA(\Delta t)$, $EL(\Delta t)$, and $EG(\Delta t)$ similarly, we find that:

$$\begin{aligned}
WTA(\Delta t) &= \left(we^{\eta_t+\eta_c}\Delta t^{1-\beta_t+\gamma_t}\right)^{1/(1-\beta_c-\gamma_c)} \\
EL(\Delta t) &= \left(we^{\eta_t-\eta_c}\Delta t^{1-\beta_t+\gamma_t}\right)^{1/(1-\beta_c+\gamma_c)} \\
EG(\Delta t) &= \left(we^{-\eta_t+\eta_c}\Delta t^{1-\beta_t-\gamma_t}\right)^{1/(1-\beta_c-\gamma_c)}
\end{aligned}$$

If the value function for cost bends more than the value function for time, i.e. if $(1-\beta_t-\gamma_t)/(1-\beta_c+\gamma_c) > 1$, $(1-\beta_t+\gamma_t)/(1-\beta_c-\gamma_c) > 1$, $(1-\beta_t+\gamma_t)/(1-\beta_c+\gamma_c) > 1$, and $(1-\beta_t-\gamma_t)/(1-\beta_c-\gamma_c) > 1$, the value per minute of the time change is increasing in Δt for all four measures. This implies that if we estimate the VTT by one of the four measures (or a combination), we would observe a VTT increasing in the size of the time change, even if the common reference-free value of time, w, was constant.

3 Data

Our data stem from a Norwegian survey conducted to establish values of travel time, variability, and traffic safety to be used in welfare-economic evaluations of transport infrastructure policies (Samstad et al., 2010). The respondents were recruited from a representative panel, and the survey was carried out on the Internet.

The survey covered both car trips, public transport (PT) trips and plane trips. In our analysis, we consider five combinations of transport mode and distance, which we analyse separately:

- Car short - car trips less than 100 km
- PT short - public transport trips less than 100 km
- Car long - car trips longer than 100 km
- PT long - public transport trips longer than 100 km

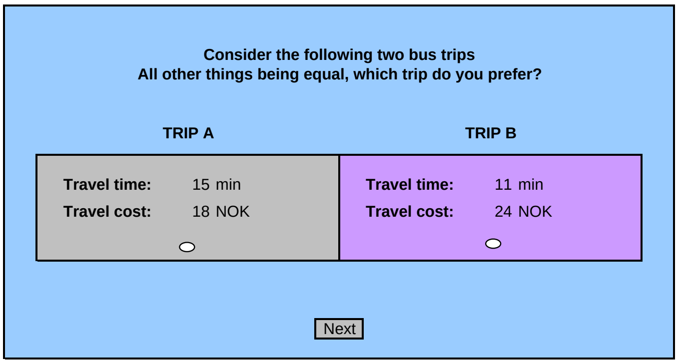

Figure 1: Illustration of choice

- Air - domestic plane trips

The survey contained several stated preference experiments, of which we use one: The choice experiment consists of nine binary choices between travel alternatives that differ with respect to cost and travel time, as illustrated in Figure 1. Always, one alternative is faster and more expensive than the other. The time and cost attributes are pivoted around the travel time (t_0) and cost (c_0) of a reference trip which the respondents reported at the beginning of the survey. The reference trip is a one-way domestic trip for private purpose, carried out within the last week (for short distance segments) or within the last month (for long distance segments). Travel time is defined as in-vehicle time without stops, except for air travellers, where travel time is measured from airport to airport. The choices are of the types shown in Tables 1 and 2. Eight of the nine choices are reference-based (two WTP choices, two WTA choices, two EG choices, two EL choices), and one choice is non-reference-based (either EG-nrb or EL-nrb).

In our analysis, we exclude respondents who answered side-lexographically (always chose left or right alternative), dropped out during the survey, or gave unrealistic reference values.[6] We also exclude air travellers with a reference travel time less than 80 minutes, because of an error in the questionnaire. These exclusions correspond to 7-9% of the observations for the car short, car long and PT long segments, and around 16-18% of the observations for air and PT short. Moreover, data are sparse for high values of reference time and cost, so we restrict our analysis to the following samples:[7]

- Car short: Cost $\leq$ 250 NOK, time $\leq$ 90 minutes.
- PT short: Cost $\leq$ 100 NOK, time $\leq$ 90 minutes.
- Car/PT long: Cost $\leq$ 1500 NOK, time $\leq$ 900 minutes.
- Air: Cost $\leq$ 5000 NOK, time $\leq$ 600 minutes, distance $\leq$ 3000 km.

[6] Unrealistic values are average speeds above 100 km per hour for land modes, average speeds above 1000 km per hour for air, costs less than 50 NOK for long distance modes, cost per kilometre less than 0.2 NOK or higher than 11 NOK for car modes.

[7] 1 NOK $\approx$ 0.12 Euro.

Table 4: Samples

Segment	Individuals	Obs	Reference-based obs
Car short	3019	27163	24144
PT short	547	4923	4376
Car long	1130	10169	9039
PT long	940	8460	7520
Air	758	6822	6064

Table 4 lists the resulting sample sizes. The sample is close to being balanced, with only 5 individuals (in the car segments) missing a few observations each. As we explain in section 4.3, our parametric analysis uses only a subsample, trimming data at 5% and 95% quantiles of Δt and Δc, which causes the samples to become more unbalanced. Table 8 in the Appendix provides summary information of the subsample used in our parametric analysis.

4 Analysis

4.1 Semi-parametric model validation

As a check of the parametric model in eq. (5), we estimate the equi-probability curves in the data and compare to those of the model. We do this separately for each data segment and choice type. To estimate the choice probabilities $P(y=1|\Delta t,\Delta c)$ as function of Δt and Δc, we use the semi-parametric framework from Fosgerau (2007), which is based on Fan et al. (1995): Let $\{(y_i,\Delta t_i,\Delta c_i)\}_{i=1}^N$ denote the sample of interest, and let Γ be the CDF of the standard logistic distribution. For a given point $(\Delta t,\Delta c)$, the choice probability $P(y=1|\Delta t,\Delta c)$ is estimated by the Local Logit Kernel estimator $\Gamma(\hat{\alpha}_0)$, where

$$(\hat{\alpha}_0,\hat{\alpha}_t,\hat{\alpha}_c) = \arg\max_{(\alpha_0,\alpha_t,\alpha_c)} \sum_{i=1}^N K_h(\Delta t_i-\Delta t,\Delta c_i-\Delta c)\log P_i(\alpha_0,\alpha_t,\alpha_c), \tag{7}$$

P_i is the logit choice probability

$$\begin{aligned} P_i(\alpha_0,\alpha_t,\alpha_c) = {} & (\Gamma(\alpha_0+\alpha_t(\Delta t_i-\Delta t)+\alpha_c(\Delta c_i-\Delta c)))^{y_i} \\ & \cdot (1-\Gamma(\alpha_0+\alpha_t(\Delta t_i-\Delta t)+\alpha_c(\Delta c_i-\Delta c)))^{1-y_i}, \end{aligned}$$

and $K_h(\cdot,\cdot)$ is a two-dimensional kernel with bandwidth h.

The estimations are carried out in Ox (Doornik, 2001), using a triangular kernel and manually chosen bandwidths. In areas where the data are sparse, the bandwidth is increased to ensure that at least 15 observations are used in each local estimation. For computational convenience, we use the same bandwidths in both time and cost dimensions.

4.2 Parametric model estimation

We estimate the parameters in our model using maximum likelihood mixed logit estimation of eq. (5): The error term ε is assumed to be logistic with mean zero and scale parameter μ (inversely proportional to the standard deviation). Log w is assumed to be individual-specific and to have a Normal distribution in the population, with standard deviation σ. This allows for unobserved heterogeneity in the VTT (note that we do not control for any observed heterogeneity, as no explanatory variables are included). We estimate a model (MXL1) with γ_t, γ_c fixed to zero, and another (MXL2) with γ_t, γ_c being free parameters. In the restricted model (MXL1), the value functions have the same curvature for gains and losses, so the entire gain-loss discrepancy is captured by the difference in levels (the η's). As a robustness check, we also estimate plain logit models, where $\log w$ is assumed to be constant.

We estimate a separate set of parameter values for each of the five data segments. Estimations are carried out in Biogeme (Bierlaire, 2003, 2005), using 500 Halton draws to simulate the individual-specific effect (see e.g. Train, 2003, for a definition).

4.3 Results

4.3.1 Semi-parametric analysis

We first regress y on Δt and Δc (as described in section 4.1). The distributions of Δt and Δc in the data have rather long right tails, implying that estimates of $P(y = 1|\Delta t, \Delta c)$ will be very unreliable for high values of Δt and Δc. Initially, we therefore only use observations where Δt and Δc are below their 90% quantiles. Figure 2 shows the estimated equi-probability curves for the car short segment, depicted in $(\log \Delta t, \log \Delta c)$-space. The bandwidth is chosen manually by graphical inspection of the estimates: Our criterion is to find the smallest possible bandwidth yielding smooth, non-decreasing and non-backbending equi-probability curves. For the car short segment, we find that a bandwidth of 0.10 is suitable (for interpretation, note that the unit of Δt and Δc are minutes and NOK, respectively).

Second, we regress y directly on $\log \Delta t$ and $\log \Delta c$. This does not produce identical results, because regressing in log space corresponds to applying smaller bandwidths for low values of Δt and Δc and higher bandwidths for higher values. Regression in log space therefore yields more uncertain estimates in the low range of Δt and Δc. To account for this, we trim data both from below (at the 5% quantiles) and from above (at the 95% quantiles). Figure 3 shows the results for the car short segment, where we find that a bandwidth of 0.15 is suitable.

As shown, the equi-probability curves for the car short segment are roughly linear, in the sense that they do not deviate systematically from linearity, except in the upper left and lower right corners where data are sparse. We find similar results for the long distance segments (not shown here): the curves are roughly linear, again excepting the upper left and lower right corners. For PT short (not shown), the pattern is less clear: Curves are not as close to linear as for the other segments, but on the other hand it is hard to find a systematic deviation from linearity. Overall, we conclude that data between the 5% and 95% quantiles do not reject the parametric model in eq. (5).

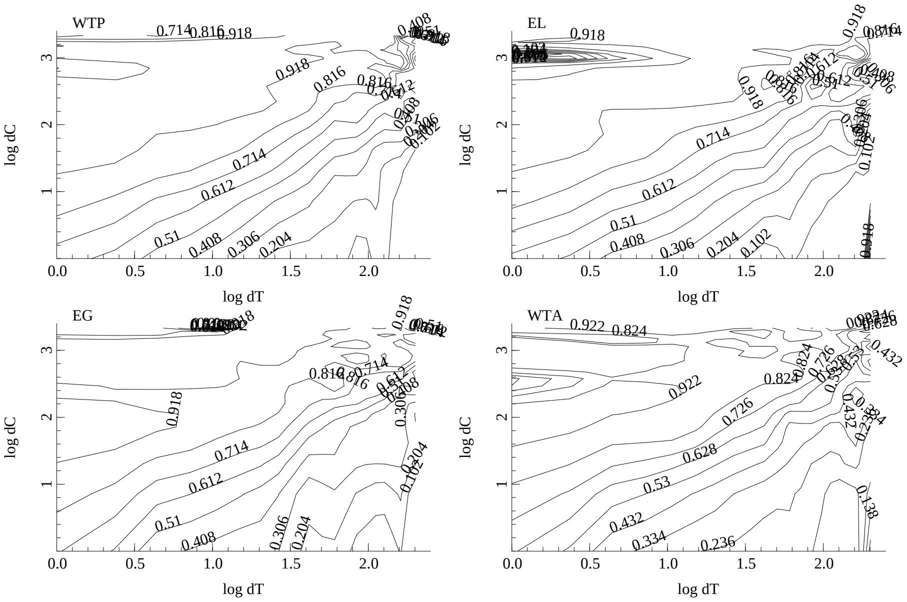

Figure 2: Equi-probability curves (local logit estimates), estimated on $(\Delta t, \Delta c)$. Car short, excluding top 10% in both dimensions. The figures along the curves denote probability levels.

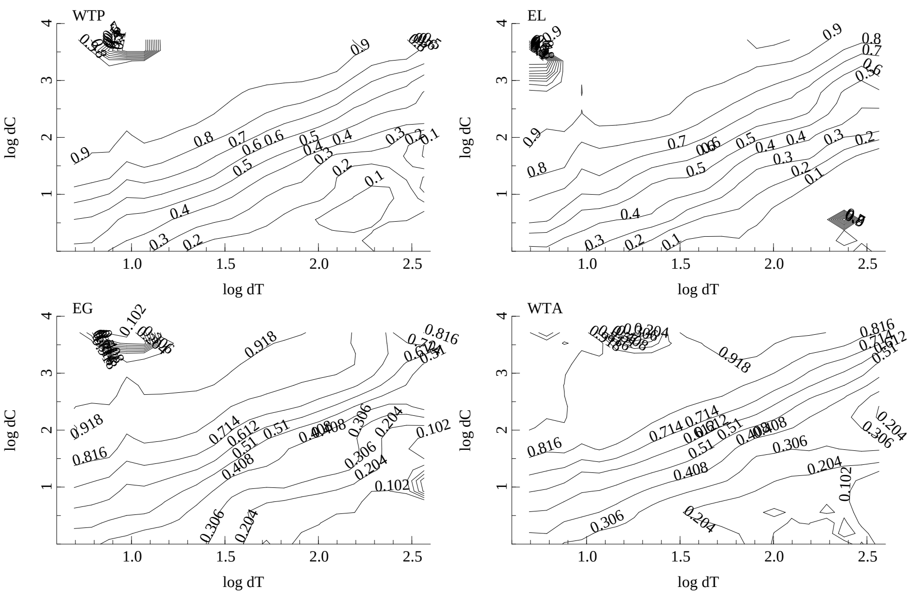

Figure 3: Equi-probability curves (local logit estimates), estimated on $(\log \Delta t, \log \Delta c)$. Car short, excluding top 5% and bottom 5% in both dimensions. The figures along the curves denote probability levels.

Table 5: Estimation Summary – Mixed Logit models (MXL1). Parameter estimates with robust standard errors in parentheses.

	Car short	PT short	Car long	PT long	Air
mean ($\log w$)	-0.46^*	-0.64^*	-0.22	-0.27^*	0.10
	(0.03)	(0.08)	(0.12)	(0.11)	(0.17)
β_c	0.19^*	0.03	0.21^*	0.15^*	0.16^*
	(0.02)	(0.08)	(0.04)	(0.04)	(0.05)
β_t	-0.02	-0.13^*	-0.03	-0.04	-0.06
	(0.02)	(0.06)	(0.03)	(0.03)	(0.04)
η_c	0.05^*	0.15^*	0.09^*	0.07^*	-0.01
	(0.01)	(0.03)	(0.01)	(0.01)	(0.01)
η_t	0.06^*	0.05^*	0.09^*	0.05^*	0.03^*
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)
σ	0.78^*	0.83^*	0.69^*	0.64^*	0.71^*
	(0.03)	(0.08)	(0.04)	(0.03)	(0.05)
μ	3.07^*	2.75^*	2.91^*	3.44^*	3.30^*
	(0.10)	(0.23)	(0.16)	(0.18)	(0.21)
Log likelihood value	-9563.1	-1681.3	-3705.8	-2999.8	-2406.2
Number of est. parameters	7	7	7	7	7
Number of obs.	23892	4375	8514	7023	5739

* denotes significance at the 5% level.

4.3.2 Parametric analysis

Based on the semi-parametric results, we limit the analysis to data between the 5% and 95% quantiles. Tables 5 and 6 present the parameter estimates. The MXL1 and MXL2 models yield practically identical value functions, so we only show the estimated value functions for the MXL2 models (Figures 4 – 6). The plain logit estimates are very similar to the mixed logit results (see Tables 9 and 10 in the Appendix), except for PT long, where the value function for cost bends more for the mixed logit model than for the logit model.

There is some variation in estimates between segments. Roughly speaking, the pattern seems to be that β_c and γ_c are significantly positive (5% level), β_t and γ_t are not significantly different from zero, and η_c and η_t are significantly positive in MXL1, but tend to become insignificant in MXL2.

From Figures 4 – 6 we see that the estimated value functions are decreasing, and that they appear to be close to piece-wise linear in the considered ranges (i.e. close to linear in the gain domain and close to linear in the loss domain). Though it appears close to piece-wise linear, the value function for cost exhibits diminishing sensitivity with respect to both gains and losses for all segments except PT short. This is significant in the sense that we can reject linearity of the value functions in both gain and loss domains (LR tests, 5% level, cf. Table 11 in the Appendix).

Table 6: Estimation Summary – Mixed Logit models (MXL2). Parameter estimates with robust standard errors in parentheses.

	Car short	PT short	Car long	PT long	Air
mean $(\log w)$	-0.46^*	-0.64^*	-0.20	-0.28^*	0.10
	(0.03)	(0.08)	(0.13)	(0.11)	(0.17)
β_c	0.19^*	0.02	0.20^*	0.15^*	0.16^*
	(0.02)	(0.08)	(0.04)	(0.04)	(0.05)
β_t	-0.02	-0.13^*	-0.03	-0.04	-0.06
	(0.02)	(0.06)	(0.03)	(0.03)	(0.04)
η_c	0.01	0.06	-0.15^*	-0.13	-0.02
	(0.01)	(0.05)	(0.07)	(0.07)	(0.10)
η_t	0.07^*	0.10	0.02	-0.06	0.03
	(0.02)	(0.06)	(0.07)	(0.07)	(0.10)
γ_c	0.02	0.05	0.05	0.05	0.00
	(0.01)	(0.03)	(0.02)	(0.02)	(0.02)
γ_t	-0.01	-0.03	0.02	0.03	0.00
	(0.01)	(0.03)	(0.02)	(0.02)	(0.03)
σ	0.78^*	0.83^*	0.69^*	0.64^*	0.71^*
	(0.03)	(0.08)	(0.04)	(0.03)	(0.05)
μ	3.07^*	2.74^*	2.88^*	3.41^*	3.30^*
	(0.10)	(0.24)	(0.16)	(0.18)	(0.21)
Log likelihood value	-9558.5	-1678.2	-3698.3	-2993.4	-2406.2
Number of est. parameters	9	9	9	9	9
Number of obs.	23892	4375	8514	7023	5739

* denotes significance at the 5% level.

For PT short, the value function for cost does not exhibit diminishing sensitivity for losses, but is not significantly different from linear in this domain (LR test, 5% level, cf. Table 11).

The value function for time does not exhibit diminishing sensitivity in either direction. However, it is generally not significantly different from linear in neither gain nor loss domain (LR tests, 5% level, cf. Table 11), the exception being PT long (loss domain), where the difference is marginally significant, and PT short (gain domain).

For the short distance segments, we have loss aversion (defined as $v_t(-|t|) < |v_t(|t|)|$ and $v_c(-|c|) < |v_c(|c|)|$) for the considered ranges of both time and cost. Loss aversion is significant in the sense that LR tests of the hypotheses of no gain-loss asymmetry in the time dimension ($v_t(-|t|) = |v_t(|t|)|$ for all t, corresponding to $\eta_t = \gamma_t = 0$) and no gain-loss asymmetry in the cost dimension ($v_c(-|c|) = |v_c(|c|)|$ for all c, corresponding to $\eta_c = \gamma_c = 0$) are both rejected at the 5% level, cf. Table 11. For the car long and PT long segments, we have loss aversion in the time dimension for the considered range of time changes, and loss aversion in the cost dimension, for cost changes larger than 15 NOK. Again the gain-loss asymmetry is significant in both dimensions (LR tests of the hypotheses of no asymmetry are rejected at the 5% level, cf. Table 11).

For air, we have loss aversion in the time dimension for the considered range of time changes, but the gain-loss asymmetry is only significant at the 10% level (cf. Table 11). We do not observe loss aversion in the cost dimension, where gains are valued higher than losses for all cost changes. Here, however, the gain-loss asymmetry is not significant (the LR test of the hypothesis of no asymmetry cannot be rejected, cf. Table 11).

Overall, these results are consistent with prospect theory: With few exceptions, the estimated value functions either exhibit loss aversion and diminishing sensitivity for gains and losses, or do not deviate significantly from this.

Moreover, the results support de Borger and Fosgerau (2008)'s proposed explanation of the positive relation between the VTT and the size of the time change, since we have $(1-\beta_t-\gamma_t)/(1-\beta_c+\gamma_c) > 1$, $(1-\beta_t+\gamma_t)/(1-\beta_c-\gamma_c) > 1$, $(1-\beta_t+\gamma_t)/(1-\beta_c+\gamma_c) > 1$, and $(1-\beta_t-\gamma_t)/(1-\beta_c-\gamma_c) > 1$.[8] Hence the value function for cost "bends" more than the value function for time, i.e. there is stronger diminishing sensitivity for money than for travel time. This implies that we would observe a value of travel time increasing in the size of the time change, if we did not take reference-dependence into account.

As a final check, we compare our results to those of de Borger and Fosgerau (2008). In Table 7, we compute the parameters $p_5 = \frac{\gamma_t}{1-\beta_t}$, $p_6 = \frac{\eta_c}{1-\beta_t}$, $p_7 = \frac{1-\beta_c}{1-\beta_t}$, and $p_8 = \frac{\gamma_c}{1-\beta_t}$, which correspond to the estimated parameters in de Borger and Fosgerau (2008).[9] The results from MXL1 should be compared to their M3R (γ_t, γ_c fixed to zero), and the results from MXL2 should be compared to their M4R.

De Borger and Fosgerau (2008) find p_5 to be significantly positive, while our estimate is never significantly different from zero. For the short distance segments, p_6 is comparable in size and sign, though not significantly positive in the MXL2 models. For the long distance segments, our estimates of p_6 differ from those of de Borger and Fosgerau (2008): For car and PT, we

[8] This is also the case for the plain logit estimates.

[9] We cannot compare our estimate of η_t directly, since we apply a slightly different model: de Borger and Fosgerau (2008) have w inside the value function for time in eq. (1).

Table 7: Comparison to de Borger and Fosgerau (2008)'s results. MXL1 results should be compared to their M3R, and MXL2 results to their M4R.

Segment	Model	$p_5 = \frac{\gamma_t}{1-\beta_t}$	$p_6 = \frac{\eta_c}{1-\beta_t}$	$p_7 = \frac{1-\beta_c}{1-\beta_t}$	$p_8 = \frac{\gamma_c}{1-\beta_t}$
Car short	MXL1		0.05*	0.80*	
PT short	MXL1		0.13*	0.86*	
Car long	MXL1		0.09*	0.77*	
PT long	MXL1		0.07*	0.82*	
Air	MXL1		−0.01	0.79*	
De Borger and Fosgerau	M3R		0.15*	0.70*	
Car short	MXL2	−0.01	0.01	0.80*	0.02*
PT short	MXL2	−0.03	0.05	0.86*	0.05*
Car long	MXL2	0.02	−0.14*	0.78*	0.05*
PT long	MXL2	0.03	−0.13	0.82*	0.05*
Air	MXL2	0.00	−0.01	0.79*	0.00
De Borger and Fosgerau	M4R	0.035*	0.09*	0.70*	0.044*

* denotes significance at the 5% level. For our results, significance tests are based on the Delta method

find that p_6 is positive in MXL1, and negative in MXL2, while for air p_6 is never significantly different from zero. The variable p_7 is comparable in size and sign (for our results, all 95% confidence intervals are within [0.72, 0.95]), while p_8 is comparable in size and sign for all segments except air. Overall, our results for the short distance segments are in agreement with de Borger and Fosgerau (2008)'s results, while our results for the long distance segments agree to some degree.[10]

[10]For comparison, de Borger and Fosgerau (2008)'s sample consists of both short and long car trips, with a large majority of trips being shorter than 100 km.

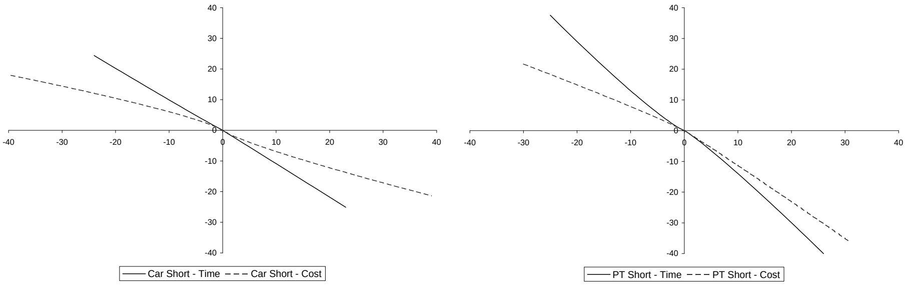

Figure 4: Value functions for car short and PT short. Value functions are depicted for the range where they are supported by the data.

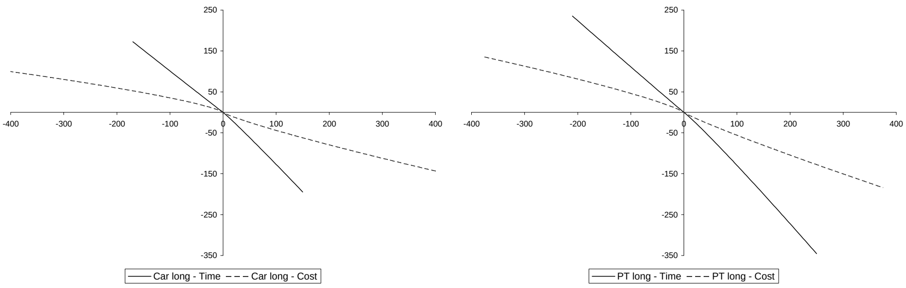

Figure 5: Value functions for car long and PT long. Value functions are depicted for the range where they are supported by the data (except for car long - cost, which has wider support)

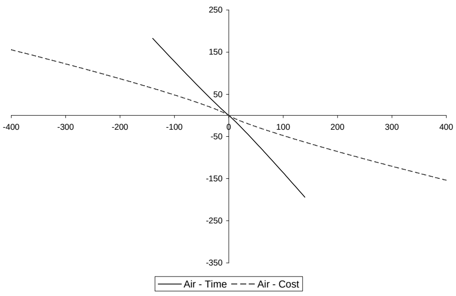

Figure 6: Value functions for air. Value functions are depicted for the range where they are supported by the data (except for cost, which has wider support)

5 Conclusion

The current paper extends the analysis in de Borger and Fosgerau (2008) and presents an empirical test with potential to falsify their proposed explanation to the phenomenon of VTT increasing with the size of the time change: That respondents have reference-dependent preferences that exhibit diminishing sensitivity for gains and losses, with a stronger degree of diminishing sensitivity for money than for travel time.

We used stated preference data with trade-offs between travel time and money that provide identification of the degrees of diminishing sensitivity for time and money gains and losses. Based on the modelling framework in de Borger and Fosgerau (2008) we formulated a parametric discrete choice model, in which choice depends on a reference-free value of travel time and reference-dependent value functions for time and money. The functional form of the value functions allows, but is not restricted to, loss aversion and diminishing sensitivity for gains and losses.

As a test of the fit of the parametric model, we compared its predicted equi-probability curves to those of the data, estimated using a semi-parametric local logit estimator. Based on this comparison, we concluded that our data do not reject the parametric model.

We estimated the value functions from our parametric model using mixed logit estimation. The results vary somewhat between the five considered data segments, but the overall picture is consistent with prospect theory: In general, the value functions exhibit loss aversion for both travel time and cost (in the time dimension we have loss aversion for the entire range of considered time changes, while in the cost dimension we only have loss aversion for part of the range of considered cost changes), the value function for cost exhibits diminishing sensitivity for both gains and losses, and the value function for time exhibits constant sensitivity for both gains and losses. We found stronger diminishing sensitivity for money than for travel time, consistent with prospect theory as the explanation of the positive relation between the VTT and the size of the time change.

References

Avineri, E., Bovy, P., 2008. Identification of parameters for a prospect theory model for travel choice analysis. Transportation Research Record 2082, 141–147.

Bierlaire, M., 2003. Biogeme: A free package for the estimation of discrete choice models. Proceedings of the 3rd Swiss Transportation Research Conference. www.strc.ch/Paper/bierlaire.pdf.

Bierlaire, M., 2005. An introduction to biogeme (version 1.4). http://transp-or2.epfl.ch/biogeme/doc/tutorial.pdf.

Burge, P., Rohr, C., Vuk, G., Bates, J., 2004. Review of international experience in VOT study design. European Transport Conference 2004.

Cantillo, V., Heydecker, B., de Dios Ortúzar, J., 2006. A discrete choice model incorporating thresholds for perception in attribute values. Transportation Research Part B 40, 807–825.

de Borger, B., Fosgerau, M., 2008. The trade-off between money and travel time: A test of the theory of reference-dependent preferences. Journal of Urban Economics 64, 101–115.

de Jong, G., Tseng, Y., Kouwenhoven, M., Verhoef, E., Bates, J., 2007. The value of travel time and travel time reliability: Survey design. Final report. Report to The Netherlands Ministry of Transport, Public Works and Water Management.

Doornik, J. A., 2001. Ox: An Object-Oriented Matrix Language. Timberlake Consultants Press London.

Fan, J., Heckman, N. E., Wand, M. P., 1995. Local polynomial kernel regression for generalized linear models and quasi-likelihood functions. Journal of the American Statistical Association 90, 141–150.

Fosgerau, M., 2007. Using nonparametrics to specify a model to measure the value of travel time. Transportation Research, Part A 41, 842–856.

Fosgerau, M., Hjorth, K., Lyk-Jensen, S. V., 2007. The Danish Value of Time Study: Final report. Report for the Danish Ministry of Transport, Danish Transport Research Institute.

Gunn, H., 2001. Spatial and temporal transferability of relationships between travel demand, trip cost and travel time. Transportation Research, Part E 37, 163–189.

Hultkrantz, L., Mortazavi, R., 2001. Anomalies in the value of travel-time changes. Journal of Transport Economics and Policy 35, 285–300.

Kahneman, D., Tversky, A., 1979. Prospect theory: An analysis of decision under risk. Econometrica 47, 263–291.

Mackie, P., Fowkes, A., Wardman, M., Whelan, G., Nellthorp, J., Bates, J., 2003. Values of travel time savings in the UK – Report to Department for Transport. Institute for Transport Studies, University of Leeds, in association with John Bates Services.

Mackie, P., Jara-Díaz, S., Fowkes, A., 2001. The value of travel time savings in evaluation. Transportation Research, Part E 37, 91–106.

Ramjerdi, F., Flügel, S., Samstad, H., Killi, M., 2010. Den norske verdsettingsstudien, Tid. Report for the Norwegian Ministry of Transport, Institute of Transport Economics, Oslo, Report no. 1053B/2010.

Samstad, H., Veisten, K., Flügel, S., Ramjerdi, F., 2010. Den norske verdsettingsstudien, Databeskrivelse. Report for the Norwegian Ministry of Transport, Institute of Transport Economics, Oslo, Report no. 1053A/2010.

Stott, H., 2006. Cumulative prospect theory's functional menagerie. Journal of Risk and Uncertainty 32, 101–130.

Train, K., 2003. Discrete Choice Methods with Simulation. Cambridge University Press, http://elsa.berkeley.edu/books/choice2.html.

Tversky, A., Kahneman, D., 1991. Loss aversion in riskless choice: A reference-dependent model. The Quarterly Journal of Economics 106 (4), 1039–1061.

Van de Kaa, E., 2005. Heuristic judgment, prospect theory and stated preference surveys aimed to elicit the value of travel time. European Transport Conference 2005.

Van de Kaa, E., 2008. Extended prospect theory – findings on choice behaviour from economics and the behavioural sciences and their relevance for travel behaviour. Doctoral thesis, TU Delft.

Wakker, P., 2010. Prospect theory for risk and ambiguity. Cambridge University Press.

Appendix

Table 8: Summary statistics of the sample applied in the parametric analysis (trimmed at 5% and 95% quantiles of Δt and Δc)

	Car short	PT short	Car long	PT long	Air
Sample size					
- individuals	3016	547	1128	939	756
- obs	23892	4375	8514	7023	5739
Reference travel time, t_0					
- min	10.0	10.0	60.0	60.0	80.0
- mean	23.4	27.3	164.8	237.0	181.2
- max	90.0	90.0	645.0	900.0	600.0
Reference cost, c_0					
- min	8.0	10.0	70.0	50.0	150.0
- mean	42.1	30.8	393.5	283.4	1144.3
- max	250.0	100.0	1464.0	1500.0	5000.0
Time attributes, t_j					
- min	-23.0	-25.0	-169.0	-210.0	-143.0
- mean	0.0	0.0	0.2	-0.2	-0.2
- max	24.0	26.0	152.0	252.0	142.0
Time attributes, t_j (gains)					
- min	-23.0	-25.0	-169.0	-210.0	-143.0
- mean	-4.8	-5.6	-33.6	-48.9	-37.6
- max	-1.0	-1.0	-9.0	-9.0	-12.0
Time attributes, t_j (losses)					
- min	2.0	2.0	9.0	9.0	12.0
- mean	4.8	5.6	34.3	47.9	37.3
- max	24.0	26.0	152.0	252.0	142.0
Cost attributes, c_j					
- min	-41.0	-30.0	-455.0	-375.0	-605.0
- mean	0.1	0.3	4.9	4.3	1.7
- max	41.0	31.0	463.0	377.0	604.0
Cost attributes, c_j (gains)					
- min	-41.0	-30.0	-455.0	-375.0	-605.0
- mean	-9.0	-8.5	-119.9	-98.2	-196.0
- max	-1.0	-1.0	-11.0	-11.0	-33.0
Cost attributes, c_j (losses)					
- min	1.0	1.0	11.0	11.0	33.0
- mean	9.8	10.4	146.5	121.6	209.7
- max	41.0	31.0	463.0	377.0	604.0
Choice variable (y)					
- min	0.0	0.0	0.0	0.0	0.0
- mean	0.7	0.7	0.6	0.6	0.7
- max	1.0	1.0	1.0	1.0	1.0

Table 9: Estimation Summary – Logit models (MNL1). Parameter estimates with robust standard errors in parentheses.

	Car short	PT short	Car long	PT long	Air
mean ($\log w$)	-0.55^*	-0.65^*	-0.20	-0.11	0.17
	(0.02)	(0.08)	(0.11)	(0.10)	(0.17)
β_c	0.24^*	0.00	0.19^*	0.10^*	0.16^*
	(0.02)	(0.09)	(0.04)	(0.04)	(0.05)
β_t	-0.03	-0.14	-0.03	-0.04	-0.04
	(0.02)	(0.08)	(0.04)	(0.04)	(0.05)
η_c	0.05^*	0.15^*	0.10^*	0.09^*	-0.02
	(0.01)	(0.03)	(0.02)	(0.02)	(0.02)
η_t	0.05^*	0.05	0.10^*	0.05^*	0.03
	(0.01)	(0.03)	(0.02)	(0.02)	(0.02)
μ	1.88^*	1.58^*	1.81^*	1.91^*	1.90^*
	(0.06)	(0.14)	(0.09)	(0.10)	(0.12)
Log likelihood value	-11867.5	-2034.9	-4347.3	-3585.9	-2965.3
Number of est. parameters	6	6	6	6	6
Number of obs.	23892	4375	8514	7023	5739

* denotes significance at the 5% level.

Table 10: Estimation Summary – Logit models (MNL2). Parameter estimates with robust standard errors in parentheses.

	Car short	PT short	Car long	PT long	Air
mean $(\log w)$	-0.55^{*}	-0.64^{*}	-0.17	-0.11	0.17
	(0.02)	(0.08)	(0.11)	(0.10)	(0.17)
β_c	0.24^{*}	-0.01	0.18^{*}	0.10^{*}	0.16^{*}
	(0.02)	(0.09)	(0.04)	(0.04)	(0.05)
β_t	-0.03	-0.14	-0.04	-0.05	-0.04
	(0.02)	(0.08)	(0.04)	(0.04)	(0.05)
η_c	0.01	0.02	-0.21^{*}	-0.27^{*}	-0.01
	(0.02)	(0.05)	(0.08)	(0.09)	(0.12)
η_t	0.08^{*}	0.10	0.05	-0.01	0.14
	(0.02)	(0.07)	(0.09)	(0.09)	(0.12)
γ_c	0.02	0.08	0.07	0.08	0.00
	(0.01)	(0.03)	(0.02)	(0.02)	(0.02)
γ_t	-0.02	-0.03	0.02	0.02	-0.03
	(0.01)	(0.04)	(0.03)	(0.02)	(0.03)
μ	1.87^{*}	1.57^{*}	1.79^{*}	1.88^{*}	1.90^{*}
	(0.06)	(0.14)	(0.09)	(0.10)	(0.12)
Log likelihood value	-11863.4	-2030.4	-4339.1	-3576.9	-2964.8
Number of est. parameters	8	8	8	8	8
Number of obs.	23892	4375	8514	7023	5739

* denotes significance at the 5% level.

Table 11: Likelihood ratio tests (p-values)

Hypothesis	p-values				
	Car short	PT short	Car long	PT long	Air
v_t linear for gains: $\beta_t = -\gamma_t$	0.19	< 0.01	0.84	0.78	0.16
v_t linear for losses: $\beta_t = \gamma_t$	0.81	0.10	0.21	0.05	0.18
v_t piecewise linear: $\beta_t = \gamma_t = 0$	0.40	0.02	0.43	0.10	0.22
v_c linear for gains: $\beta_c = -\gamma_c$	< 0.01	0.30	< 0.01	< 0.01	0.01
v_c linear for losses: $\beta_c = \gamma_c$	< 0.01	0.73	< 0.01	0.03	0.01
v_c piecewise linear: $\beta_c = \gamma_c = 0$	< 0.01	0.07	< 0.01	< 0.01	0.02
v_t and v_c piecewise linear: $\beta_t = \gamma_t = \beta_c = \gamma_c = 0$	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
No gain-loss asymmetry for time: $\eta_t = \gamma_t = 0$	< 0.01	0.01	< 0.01	< 0.01	0.06
No gain-loss asymmetry for cost: $\eta_c = \gamma_c = 0$	< 0.01	< 0.01	< 0.01	< 0.01	0.85
No gain-loss asymmetry: $\eta_t = \gamma_t = \eta_c = \gamma_c = 0$	< 0.01	< 0.01	< 0.01	< 0.01	0.18

Note: The piecewiese linear formulations have separate slopes for gains and losses.